DOCUMENT RESUME

ED 411 683 FL 024 735

AUTHOR Jenkins, Susan; Parra, Isabel
TITLE An Investigation of Interactional Authenticity in International Teaching Assistant Interview Testing.
PUB DATE 1997-03-00
NOTE 40p.; Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (31st, Orlando, FL, March 11-15, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Chinese; Discourse Analysis; *English (Second Language); *Foreign Students; Graduate Students; Higher Education; Interaction; Intercultural Communication; *Interpersonal Communication; Interviews; *Language Proficiency; Spanish; Speech Skills; *Teaching Assistants; Test Wiseness; *Verbal Tests

ABSTRACT

A study investigated how non-native speakers of English used linguistic and non-linguistic skills to manage interaction in an interview-style verbal test of English proficiency. Subjects were eight international teaching assistants, four native speakers of Spanish and four native speakers of Chinese, who took a standardized oral interview examination. Analysis of the recorded interviews indicate that those subjects who used appropriate discourse management skills and negotiated a degree of control over the interview process were able to compensate for weaker linguistic proficiency. Successful test-takers were able to modify the power dynamic and reduce the interview asymmetry. Those who framed the interview as a discussion or conversation among peers were more successful than those who framed it as an examination. In addition, evaluators tended to spend less time and energy in the interaction when they were not engaged. Contains 46 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 411 683

An investigation of interactional authenticity in international teaching assistant interview testing.

A Paper presented at Annual TESOL Convention,
Orlando, Florida March 11-15, 1997

Susan Jenkins and Isabel Parra
University of Cincinnati

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Susan Jenkins
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Address for correspondence:
Dr. Susan Jenkins
College of Education
P.O. Box 210002
University of Cincinnati
Cincinnati, OH 45221-0002

Tel: 513-556-3590 (office)
513-231-6659 (home)

Fax: 513-556-2483

E-mail: Susan.Jenkins@UC.EDU

FL024735

BEST COPY AVAILABLE

Abstract

Under the influence of Gumperz' (1982) sociolinguistic theory of conversational inference, and the concept of frames (Goffman, 1974; Tannen, 1993), this study explored the contribution of linguistic and non-linguistic contextualization cues in promoting interaction and involvement in oral interview test situations. The participants were four Spanish and four Chinese international teaching assistants (ITAs) who speak English as a second language. Microanalysis of videotaped tests indicated that interviewees who exhibited appropriate discourse management skills, and negotiated a degree of control over the interview process were able to compensate for weaker linguistic proficiency. Successful test takers in this study were able to modify the power dynamic and to reduce the interview asymmetry. Those who framed the interview as a discussion or conversation among peers, were more successful than those who framed it as an examination. Evaluators tended to spend less time and energy in the interview interaction when they were not engaged.

"The walls could have been talking to her":

Interaction and involvement in an oral proficiency interview.

The expanding global economy has emphasized the need for multi-lingualism among the peoples of the world. As people cross borders and languages to work and study, we have come to rely increasingly on a variety of gatekeeping tests (Erickson & Schultz, 1982) of oral language proficiency to determine fitness for the projected language communication task.

This paper is concerned with oral proficiency assessment of one group in this trend towards internationalism: international teaching assistants (ITAs) in North American universities. The presence of non-native English speaking TAs in the classroom has led to a perception of a "foreign TA problem" (Bailey, 1984), which in turn has prompted legislation in a number of states to mandate oral proficiency assessment (Thomas & Monoson, 1993). Because of the potentially discriminatory aspects of singling out one group for testing, it is important to investigate the multivariate nature of language proficiency as it interacts with the test method and learner characteristics (Bachman 1990). Such studies will help us to provide a comprehensive picture of ITA performance, especially in terms of predictive validity.

Author (1996) analyzed videotaped oral proficiency tests of four Spanish and four Chinese speakers, and found that problematic linguistic cues such as pronunciation, stress, intonation and rhythm, pausing and discourse markers were strongly correlated with low failing and borderline passing scores. For the borderline cases, discourse and conversational management strategies, and the ability to reduce the interview asymmetry became the most salient criteria in evaluator ratings, determining passing or failing decisions. This present study investigates the nature of the linguistic and nonverbal contextualization cues that contributed to the evaluators' assessments of communicative competence. We argue that the ITAs' interpretations of their role and status, and the interviewer-interviewee co-construction of meaning through negotiation and discourse strategies played a significant role in the evaluations of the teaching assistants

Theoretical Framework

This study is influenced by Gumperz' (1982, 1992) sociolinguistic theory of conversational inference, and by the concept of frames (Goffman, 1974, 1981; Tannen, 1993).

Gumperz' believes that face-to-face encounters entail the co-construction of meaning as participants interpret the on-going linguistic, semantic and pragmatic features. Interlocutors react to these "contextualization cues" to infer each other's intentions. Tannen (1989, p. 11) used the term "conversational styles" to describe the same phenomenon. Situated interpretation is also influenced by "participants' personal background knowledge and their attitudes toward each other, sociocultural assumptions concerning role and status relationships as well as social values associated with various message components" (Gumperz, 1982a, p. 153).

Interpersonal involvement is necessary "to attract and sustain others' attention" (Gumperz, 1982, p. 4). The notion of involvement is key to Gumperz' definition of understanding in interpersonal interaction. Tannen (1989) explored the use of the term in the work of several scholars of spoken and written language production, and concluded that, although actual terminology may vary, the majority concur that involvement, and thus understanding in interpersonal encounters entails that participants recognize and respond with rapport to the "familiar patterns" (p. 13) in the discourse. Like Gumperz, Tannen (1989) sees involvement as "not a given but an achievement in conversational interaction" (p. 12).

Gumperz' approach is complemented by Goffman's frame analysis (Goffman, 1874, 1981) by which participants choose the appropriate frame, or alignment we take up to ourselves and the others present as expressed in the way we manage the production or reception of an utterance" (Goffman, 1981, p. 128). Frames function as organizational guidelines for the events in which people are involved. Face-to-face interpersonal involvement requires the interlocutors to signal and interpret cues that will correctly identify the appropriate key, or footing for the interaction. Miscommunication may occur when participants' socio-cultural or linguistic backgrounds differ. In an oral proficiency test this possibility may be magnified, as evaluators expect familiar patterns, while test takers may find their skills diminished by nervousness.

Issues in Language Testing

Bachman (1991) claimed that "one legacy of the 1980s is that we now know that a language test score cannot be interpreted simplistically as an indicator of the particular language ability we want to measure" (p. 677). Bachman's (1990; 1991) model of language testing recognizes the multivariate nature of language ability. Language knowledge consists of linguistic and pragmatic knowledge that together make up the contextualization cues available to speakers and listeners. Thus, tests which measure language ability should ensure both situational authenticity (in which the test method matches the language use situation), and interactional authenticity (the extent to which language ability is engaged in completing the task).

Three types of test, according to Hoekje and Linnell (1994), have come to the fore: The Test of Spoken English/Spoken Proficiency English Assessment Kit (TSE/SPEAK) (Educational Testing Service, 1990), the ACTFL/ETS Oral Proficiency Interview (American Council on the Teaching of Foreign Languages, 1986) and performance tests, in which tasks related to the context of teaching are evaluated. Hoekje and Linnell measure these tests against the standards proposed by Bachman, and conclude that situational and interactional authenticity are most effectively demonstrated in teaching performance tests.

Nevertheless, decisions about which test to use are often influenced as much by practical realities as by factors such as situational and interactional authenticity. The TSE/SPEAK is often the test of choice because it is easy to administer, although the test taker speaks only to a voice on an audiotape. The ACTFL/OPI has especially been attacked as lacking authenticity, reliability and validity (Bachman, 1988; Bachman & Palmer, 1983; Bachman & Savignon, 1986; Lantolf & Frawley, 1988). Performance tests, although high on situational and interactional authenticity, are risky in the sense that they test non-language factors overtly, namely teaching skills. In most cases, native English speaking TAs are never subject to such tests, thus raising the specter of potential litigation (Brown, Fishman, & Jones, 1990).

In the present study, the Oral Proficiency Test designed for use at a large research-oriented university had to consider as many "political" as validity and reliability issues in

developing the test. The university and departments were concerned about equity issues and decided that ITAs would not be singled out. This decision has had two major implications for test design. First, all graduate students receiving university funding are tested. Second, a teaching performance test was considered inappropriate because many of the students taking the test would be research assistants or tuition scholarship holders uninvolved with teaching.

The test is a videotaped interview in which the candidate performs four different language tasks (described in Appendix A) before three trained evaluators, who are graduate students in language-related disciplines. Evaluation criteria focused on pronunciation, grammar, fluency, and comprehensibility. This was necessitated by the institution's insistence that only formal language factors be rated. Each evaluator rates each section of the test according to the given criteria (see Appendix B), and then all scores are averaged. The student passes the test on receiving a score of 3 on comprehensibility which is an overall impression of the intelligibility of the student in spite of problems in pronunciation, grammar and fluency.

We argue that the tasks are situationally authentic because they sample typical academic tasks required of all graduate students. The first section gives the examinee control over the discourse domain (Douglas & Selinker, 1985), and the graph and syllabus presentations are closely related to ITA duties. However, since the examinee does not control the topics and subject matter in sections 2, 3, and 4, the test is possibly subject to some of the same criticisms leveled against the ACTFL/OPI, namely that it may be "a criterion-reductive, analytically derived, norm-referenced test of how well an individual can deal with an imposition" (Lantolf & Frawley, 1988). The next section will review the literature on the assessment of oral proficiency involving face-to-face interactions, via ACTFL/OPI or performance type tests. We claim that evidence from these studies indicates that oral interviews contain features of negotiated discourse in support of Gumperz' (1982) view of jointly constructed negotiation to establish meaning. As such, these aspects of human communication have not been given the prominence they deserve in discussions of the authenticity, reliability and validity of oral proficiency interviews.

Review of Relevant Literature

Oral Proficiency Interviews

Critics have claimed that the ACTFL/OPI is based on a unitary view of language, and confounds the assessment of language proficiency with the method of assessment (Bachman, 1988). Lantolf & Frawley (1988) rejected the OPI because it is not based on an acceptable theory of human communication, which to them is a Vygotskyan perspective of a dynamic open system. The OPI fails on this account because the interviewer is in complete control of topic nomination and discourse development in a closed system.

In a similar vein, van Lier (1989) rejected the OPI as an instance of conversation. He considers that face-to-face interaction must involve decision-making opportunities for the participants, and goal-relatedness, which is akin to a component of Bachman's strategic competence (1991). For van Lier, the OPI fails as conversation because it is not characterized by "reactive and mutual contingency" that is, equal opportunity for topic nomination and turn-taking, but rather by "asymmetry and interviewer control" (pp.497-498).

Young & Milanovic (1992) found evidence of asymmetrical contingency in oral proficiency interviews, with the test taker mainly reacting to examiners' topics. The candidate spoke twice as much as the examiner, but dominance was maintained by greater reactiveness on the candidate's part, and by the interviewer's giving the candidate the floor frequently, but being able to regain it at will. Young & Milanovic refer to this as "a candidate's obligation (rather than right) to speak" (p.417). They also found that reactiveness seemed to be a function of interpersonal characteristics. This aspect was not available for further study because the data sources were audiotaped interviews, and as the authors noted, a great deal of interaction in conversational encounters occurs because of participants reactions to non-verbal cues.

Several studies suggest that proficiency assessments may be much more dependent on non-linguistic criteria than has been previously considered. Halleck (1995) investigated the syntactic basis of the OPI by comparing measures of syntactic maturity with holistic ACTL/OPI scores. Syntactic maturity generally increased as proficiency level rating rose, but no differences

were found in syntactic maturity between ratings of Advanced and Intermediate level subjects. Halleck interviewed 100 OPI raters and determined that communicative factors related to discourse management and conversational strategies were more significant to the raters than discrete-point grammatical features (Halleck, 1992, p.231).

Ross (1992), Ross and Berwick (1992) and Berwick and Ross (1993) demonstrated that non-linguistic contextualization cues play a more important role than has been previously recognized in oral interviews. Although characterized by unequal status, interviews assume the characteristics of dyadic interaction, including "foreigner talk" accommodative discourse.

Ross (1992) and Ross and Berwick (1992) examined accommodative questions in 16 audiotaped OPI tests. Accommodation occurred when evaluators perceived problems with the appropriateness of the previous response, had difficulty in understanding the response, or developed low expectations of the candidate's proficiency as the interview progressed. Accommodative questions discriminated among the proficiency interview ratings. When the frequencies of ten accommodation question types were compared with actual proficiency rating, the fit at ratings at the lower and upper ranges (1+ and 3) was good, but the intermediate ranges (2 and 2+) were less clearly differentiated. Ross & Berwick suggested that interviewers at the mid ranges may tend to over- or under-accommodate, but an alternative interpretation, in line with Halleck's (1992) findings, is that characteristics of the non-verbal and paralinguistic interaction constituted unknown variables in the evaluation. Since the data were analyzed from audio tapes, the non-verbal channel was not available for analysis.

Berwick & Ross (1993) examined accommodation patterns between Japanese and English language interviewers, and found differences that they attributed to culturally specific pragmatic and schematic interpretations of the interview speech act. Japanese interviewers provided many more accommodative and topic maintaining moves to avoid communication breakdown. These differences in "familiar patterns" affect ratings of second language interviews, and in fact it would be unrealistic to expect otherwise.

Neu (1990) noticed that in an ESL placement test, which was rated according to linguistic criteria, it was "generally acknowledged by the instructor-raters that 'fudging' of the scores occurs in order to reflect a factor not taken into account but felt to be critical to overall communicative competence: nonverbal competence in L2" (p.123). She studied videotapes of two students who received roughly equivalent scores, and hence placement in class. The verbally weaker student was able to take control and "bluff his way through" (p.135) because of his nonverbal competence. In contrast, the linguistically stronger student was perceived as less fluent than he actually was because he was unable to use appropriate nonverbal behavior.

Interactive Performance Tests

Interactive Performance tests for ITAs developed since the early 1980s have sought situational authenticity through teaching tasks performed in front of an audience or a panel. Studies of instructional discourse have revealed a variety of contexts in which ITAs need to demonstrate language proficiency (Madden & Myers, 1994). In addition, several studies have demonstrated a powerful role for effective interactive behavior in instruction. For example, Bailey (1982) identified bonding and elicitation moves as significantly correlating with rated teaching effectiveness. Nelson (1991) found that self-disclosure on the part of the ITA created a rapport with undergraduates that overcame linguistic difficulties. Inglis (1993) determined that the factor "attentiveness" accounted for most of the variance in a communicative style measure, and correlated highly with Bailey's "bonding moves." Inglis' explanation is that attentiveness is "both active and reactive" (p. 100), and is promoted when instructors pay close attention to both the verbal and non-verbal cues of their students.

Thus, as Seliger suggested in 1985, pragmatic competence may assume a more important role in communication when linguistic competence is limited. But there is also evidence of a language threshold level beneath which compensatory strategies or pragmatic skills do not make up for linguistic problems. Bailey (1982) established a threshold rating of 2 on the OPI for effective classroom functioning. Halleck and Moder (1995) determined that students who failed a teaching test had serious problems with their language skills to the point that their deficiencies

dominated in the evaluation. For ITAs in the borderline provisional pass range, on the contrary, there was little consistency in the relative contribution of language skills versus teaching skills in their evaluations, suggesting that evaluators were balancing these factors differentially.

In summary, several studies of oral proficiency interviews and performance tests have indicated that interactive involvement, linguistic and non-linguistic contextualization cues and frames are all implicated in the evaluation rating to a much greater extent than the scales, guidelines and criticisms of such tests would suggest. Although located within the frame of an interview, and thus maintaining asymmetrical relations, the literature shows that interviewers try to change the footing to incorporate conversational styles, use accommodative questions and pay attention to communicative factors. Interviewees who exhibit appropriate discourse management skills, and negotiate a degree of control over the interview process, seem to be able to compensate for weaker linguistic proficiency. The study described in this paper examines closely the contextualization cues in a test situation that promote interactional involvement between the candidate and the evaluator, and their influence on ratings of oral proficiency.

Method

Participants

Videotaped Oral English Proficiency Tests (OEPT) of four Spanish speaking and four Chinese speaking ITAs are the data for the discourse analysis. Table 1 shows their gender, country of origin, department in which they are majoring, and OEPT scores.

[Insert Table 1 about here]

Ana and Lu received high scores as "completely comprehensible." Carmen, Alejandro and Shiru received marginal passing scores, and were rated "generally comprehensible." Wang failed, but at the borderline, indicating that some of the evaluators gave him some passing scores, but overall found him to be "occasionally incomprehensible." Pedro and Ying were judged to be "generally not comprehensible."

The departments of Spanish, Physics, Mathematics and Music primarily employ TAs to teach classes, lead laboratory or recitation sessions, give tutorials, or maintain office hours. The

students in Education and Neuroscience were research assistants, and Chemical Engineering students were graders, or led recitation sections. The two Spanish Department ITAs were assigned to teach elementary Spanish to undergraduate students.

Procedures and Analysis

To triangulate the analysis of the videotapes, data was collected from a variety of sources. The evaluation reports for the eight students were examined. These reports contain the rating sheets of each evaluator in which they assign a numerical value between 1 and 4 to each of the linguistic categories (pronunciation, grammar, fluency, and comprehensibility) for each of the four tasks. Each of the three evaluators also writes comments about the candidate's performance, and gives an overall assessment of classroom readiness. One randomly selected evaluator from the team in the videotaped test was asked to view the tape again and to give "think aloud" comments (Ericsson & Simon, 1993) about their reactions and perceptions as they watched the tape. Their reactions were audiotaped and transcribed. The four students from the PRC were also invited to watch their test, to record their impressions about their performance, and recall what they could about their decisions and strategies for the test. These reports were also audiotaped and transcribed. Finally, two highly proficient Chinese TAs who worked in the OEPT program were asked to view the videos and to make comments from their own unique cultural perspectives about the performances of the PRC students. To provide base-line data of native speaker performance in the same context, two U.S. graduate students were also videotaped as they took the oral proficiency test.

The second and fourth sections of the OEPT, the conversational questions and the presentation of a syllabus, were targeted for detailed analysis. Interviews with evaluators revealed that they believed these were the most revealing in assessing proficiency. One evaluator said of the conversational questions that, "it is a test of their ability to be spontaneous in English," and another stated that students "can make it their own," that is, take control of the topic. Likewise, the evaluators felt that the syllabus was illuminating because "in this section, there is a better illustration of what their role might be like in being a TA."

The videotapes were first transcribed by a graduate student and then viewed repeatedly by the researchers independently with the transcription in hand to check for accuracy and to add descriptions of nonverbal behavior. The final transcription represented their consensus. Transcription conventions followed Tannen (1989), and are described in Appendix III.

The discourse analysis focused on the contextualization cues that promoted a sense of involvement for the evaluators, enabling them to infer the communicative competence of the ITAs. In the joint construction of meaning, the ITAs were listeners when receiving instructions for the tasks, or when listening to questions from the evaluators. They were the primary speakers when they took the floor to respond to questions or present the syllabus. Three categories emerged from the discourse analysis to illustrate successful and unsuccessful ways in which the ITAs managed these interlocutor responsibilities: 1) Listening behavior, or backchannel communication (Yngve, 1970), when the ITAs signal their active participation as good listeners while remaining in the non-primary speaker role (Clancy, Thompson, Suzuki, & Tao, 1996). Typical responses are "yeah," "mm huh," "ok," and nodding. 2) Turn-taking behavior (Sacks, Schegloff, & Jefferson, 1974) by which the listener responds to familiar cues signaling comprehension and assumption of primary speaker status. Three types were identified: acknowledging tokens (Jefferson, 1984), confirmation checks, and clarification requests (partial and complete). 3) Attempts to frame the discourse to create content knowledge expectations, and to achieve interactive involvement and rapport, once the ITA had the floor.

Results

Interview discourse is typically characterized by asymmetrical dominance (Van Lier, 1989; Young & Milanovic, 1992); thus candidates are initially limited by the interviewer's right to nominate topics and control the discourse. They are also under an obligation to adhere to Grice's cooperative principle, particularly the maxim of quantity (Grice, 1975). Nevertheless, the analysis revealed that candidates who were judged to be communicatively competent used appropriate linguistic and nonverbal contextualization cues (Gumperz, 1982), and successfully framed the context and interpretation of the discourse (Goffman, 1974, 1981; Tannen, 1993).

Listening Behavior

High scoring candidates exhibited a variety of active listenership behaviors as they received instructions or test questions. Combined with their proficient linguistic skills, this behavior was a powerful indicator for the evaluators of their "comfort" level with English. Ana and Lu were in this category. While they were listening to instructions, they maintained eye contact, nodded slightly at appropriate intervals (approximately 3-4 seconds), and employed frequent and appropriate backchannels such as "ok," "yeah, "sure," and "uh huh." Their faces were expressive of interest and involvement as they reacted physically by smiling, laughing, raising their eyebrows, or grimacing. Body position was inclined towards the evaluators.

The three students at the borderline pass level, Alejandro, Carmen and Shiru, differed considerably in their linguistic competence. Carmen's linguistic competence was high. She was clearly fluent and could control complex grammatical structures with relative ease. Her pronunciation was only occasionally problematic when her rate of speech increased. Shiru and Alejandro, on the other hand, were linguistically weaker. In fact, when one of us initially viewed the tape, she commented that Alejandro should definitely have failed the test. He had difficulty in understanding the instructions or questions on all four sections of the test. He frequently groped for vocabulary, code switched and transferred from Spanish. Shiru's listening comprehension was good, but her discourse was frequently hesitant and groping.

Alejandro's and Shiru's listening behavior contained the same features as that of Ana, Lu and the N. American TAs, as examples 1 and 2 illustrate.

[1] Eval.: Now, I'm going to ask you some questions, on topics of general interest ,

Alej.: mhm

Eval.: You give us your opinions, and remember, there's no right or wrong answers, on to these questions, we are only interested in the way you express yourself.

Alej.: OK.

[2] Eval.: Question number two.

Shiru: mhm

Eval.: What kind of customs, help young people, find their future spouses, in China?

Shiru: What kind of.. er?

Eval.: Customs.

Shiru: Customs. oh it's different,

As Alejandro listened to the instructions, he leaned forward, inclining his ear towards the evaluators, and nodding and smiling frequently. Shiru's behavior was remarkably similar.

In contrast, Carmen's listening behavior indicated attentiveness, but impatience. At the beginning of the test she maintained eye contact while listening to the instructions and responded with verbal and nonverbal backchannels. However, her nodding involved deep, rapid and repeated movements of the head as though she understood and was impatient to get going. She was unsmiling throughout the instructions, and the test itself. As the test progressed, she withdrew eye contact during the directions and sat upright with her elbows on the table, chin in hand, looking down.

The borderline failing student, Wang, exhibited listening behavior that was passive, still, and non-interactive. He maintained eye contact, and sat forward, but did not backchannel His facial expression was serious and did not change, except at one point when the evaluator told him he could use the "blackbird" instead of the "blackboard," at which point all evaluators laughed and Wang joined in belatedly. Ying also exhibited the same still body language and non-vocal listening behavior, attending to all instructions with the same slight smile. Pedro's listening behavior was interactive and often similar to Alejandro's and Shiru's. However, his linguistic competence was too low for his listenership behavior to compensate for his poor fluency.

Attentive listening behavior that employs appropriate verbal and non-verbal cues is the first indication for evaluators that the candidate is communicatively competent. Various studies in conversational analysis have demonstrated the importance of features that suggest expected sociocultural indicators of supportive listenership (Gumperz, 1982; Tannen, 1989). By listening in a reactive manner, Ana, Lu, Alejandro and Shiru, were demonstrating conversational

management skills in ways similar to North American expectations. Videotapes of two English speaking North American graduate students (NSEs) indicated that their behavior incorporated all the interactive and attentive listenership features described above as they listened to the same instructions. Non-native speakers in oral interview tests thus have an important opportunity to frame their evaluators expectations at an early point in the interview.

Turn-taking Behavior

This category is defined as those "turn-entry devices" (Sacks, Schegloff, & Jefferson 1974) that signal the taking of the floor by the candidate. Their function is to provide planning time, ratify the previous utterance, or show attentive listenership (Chafe, 1985; Tannen, 1989). They occur immediately after the evaluator has finished giving the instructions or asking a question. The data show three strategies used by the interviewees: Acknowledgment tokens (Jefferson, 1984), confirmation checks, and clarification requests.

Acknowledgment tokens

Turn-taking in conversation typically involves a variety of responses to indicate that the previous utterance has been understood and that ownership of the floor has changed. In this study, the most frequent responses were vocalizations such as "umm" followed by "yeah," "well," and "ok." Occasionally, the ITAs also used a strategy employed frequently by the native speakers, namely a sotto voce repetition of a key term, as Pedro did in example 3.

[3] Eval: OK the second question is, describe a recent situation, in which you felt frustrated.

Pedro: Recent situation...well..when I try to express myself in English, (laughs)

Eval: (laughs)

Example 4 demonstrates the power of an acknowledgment token to convey an impression of being in control:

[4] Eval: Describe the application procedures, before being admitted, to the graduate program, you are currently enrolled in.

Alej: I don't understand, excuse me.

Eval: What procedures, did you have to go through, to come to CCM?

Alej: O::K OK. I will study my master, my master in music and composition, but I was, I was conducting an orchestra, and teaching music in one university in my country, my ... my aspiration my.. my wishes, I want to improve my, my knowledge, knowledge? in music, in conduction basically in composition, I am very very interested in teaching, I like to teach music in my country, this is the principal reason, about that I am and study here.

Eval: Did you have to fill out, lots of paperwork?

Alej: Ye::s, in my country and in the United States, (laughs) all the time, many MANY papers, yes.

As he took the floor, Alejandro said "O::k," elongating the vowel to exude complete confidence, and began to answer an entirely different question, nodding smiling and interacting with the evaluators. The evaluator regained the floor and got back on track via a simple yes/no accommodated question. Again, Alejandro, employed the same vowel elongating strategy on "ye::s" as he replied. Alejandro had not understood the question, and it is likely that his attentive listening behavior of leaning forward and inclining his ear towards the evaluators was attributable to his listening comprehension problems. However, his demeanor during his inaccurate response lent credence to his confident, acknowledgment tokens.

The least proficient student, Ying, used an inappropriate strategy that reinforced perceptions of his lack of comprehension and fluency, as in example 5:

[5] Eval: OK. Question number three, what do you think, are the most important characteristics, of an effective teacher?

Ying: ...[7 second pause in which he looks down] To be a good teacher ..um ... I think ..the teacher, should understand, the situation, the students, ... eh, are.

Ying's choice of silence and lack of eye contact as a backchannel occurred throughout his test and led the evaluators to question his listening comprehension and fluency. This silence was in marked contrast to the behavior of the highly proficient NNS and the N. American TAs. They

began with filled pauses, for example, a facial grimace or an acknowledging token such as "well, let me see" muttered quietly, or repetition of the key phrase.

Highly proficient Lu and Anna also frequently used the device of repetition of key words, as did the N. American. TAs . Shiru employed repetition as a clarification request (as in example 2 above). Ying, although he often repeated key words appropriately, was not fluent enough to follow through, as shown in example 6:

[6] Eval: And is the final comprehensive?

Ying: Oh yeah, final. we we got the whole ... er text final will cover the whole thing.

Ying used the strategy of repetition as a way of grasping at the language given to him. He used appropriate falling intonation, but did not have the fluency to take advantage of the planning time he gave himself but he did not realize it was a planning strategy, and his response is not very comprehensible. Ying's extended silences in place of verbal and non-verbal backchannels also caused the evaluators to believe he lacked fluency.

Confirmation Checks

This category is defined as a response which repeats part of the previous utterance and which receives an affirming verbal or non-verbal backchannel response without elaboration. The assumption on the part of the evaluator is that the candidate has understood the question. Confirmation checks repeat key words or make use of paraphrase, and incorporate appropriate stress and yes/no question intonation. In example 7, Lu paraphrased "attraction," demonstrating not only that she comprehended but that she could supply synonyms:

[7] Eval: Please, describe for us, a famous tourist attraction, in your country, that you would recommend to a visitor.

Lu: A tourist site?

Eval: mhm

Alejandro and Shiru also used the strategy effectively, as examples 9 and 10 demonstrate.

[8] Eval: OK. describe for us, your hometown.

Alej: My hometown?

Eval: mhm

Alej: OK.

However, Ying's use of confirmation checks served to emphasize both his lack of fluency and imperfect grasp of paralinguistic and prosodic aspects of English. Ying's long pause in example 9 as he looked down and to the side allowed the evaluators to infer that he had not understood the question:

[9] Eval: OK. first question, describe a famous tourist attraction in your country, that you would recommend, to a visitor.

Ying: ... [3 second pause] In MY country?!

Eval: Yeah, if I'm going to China, then, what would you recommend that, eh I

Ying: [OK]

Eval: I do in China?

Ying: OK (laughing)

When he looked up at the evaluators, he gestured to himself and asked what may have been a confirmation check, but was interpreted as lack of comprehension. Ying's question placed stress on "my" coupled with high rise intonation on "country" to give the impression that he was surprised that anyone would be asking a question about his own country. The evaluator immediately interpreted this as lack of comprehension and made an accommodation towards simplifying the question.

A confirmation check probably plays the same role as other turn-entry devices in providing planning time for the speaker. If it occurs immediately with an appropriate intonational contour, the evaluators recognize it as a check and assume it to be a familiar and expected conversational cue.

Clarification Requests

These are realized in two ways: as a request for a partial repetition of key terms, or as a statement of non-comprehension that necessitates a complete repetition or accommodation of the question. Examples 10 and 11 illustrate partial repetition requests:

[10] Eval: OK. what do you think, is the best way to relax, during stressful times?

Alej: To relax, when?

Eval: During stressful times,

Alej: In my free time?

Eval: During stress, when you're uptight or,

Alej: OK when I feel upset or tense?

Eval: What does relax you?

Alej: We::ll, maybe to take, one hot sho ..sho..shower, shower?

Eval: uh huh

[11] Eval: OK. OK. the last question is, what do you think, is the best way to relax, during stressful times?

Pedro: During what sorry

Eval: To relax,

Pedro: During?

Eval: During stressful times,

Pedro: I don't know, I think ...go to bed, and sleep for twenty minutes, that's a way of.. take a pill .. I don't know, (laughs)

Eval: (laughs) Have a party? (laughs)

Pedro: Also.

Although on the surface, Alejandro and Pedro seemed to have the same comprehension and fluency difficulties in examples 10 and 11, the differences in the nonverbal channel were eloquent. Alejandro was in control; Pedro was not. Pedro was extremely nervous, mopping his face frequently, and responding as briefly as possible. He was clearly uncomfortable and responded tentatively, usually ending his utterances with a rising intonation. The evaluator's concern for his affective state can be seen in his nomination of the topic "party" which had been previously introduced by Pedro himself. His inappropriate response of "also" was a transfer of the Spanish "también," used to mean "yes, that too."

Example 11 shows that Alejandro did not comprehend the question, and that his clarification request strategy did not help him. His confirmation request, "in my free time?" revealed that he found only the "-ful" morpheme of "stressful" salient, and equated it with a semantically similar miscue for the context "relax." When he finally understood the accommodated question, he leaned back, smiled in triumph and swept his hand upward as he demonstrated his comprehension by paraphrasing "uptight." The evaluator's next question served to re-establish the interview context, but Alejandro was not to be drawn into the interview frame at this point. He pulled on his ear, smiled, kept eye contact, nodded and gave his brief response in a conversational frame. He obtained yet another confirmatory vocalization from the evaluator about his pronunciation of the word "shower," as he did also in example 4 for the word "knowledge." Finally, the evaluator asked him an accommodative question that he knew to be in Alejandro's discourse domain of music.

As with confirmation checks, clarification requests may give planning time, or are the result of not hearing a crucial phrase. All examples in the data have the expected prosody of rising intonation and call forth a repetition of the noun phrase queried. Thus, evaluators do not infer that partial repetition requests indicate poor comprehension, and they willingly respond to the request, indicating that they interpret these as normal discourse management features.

Complete repetition requests occurred rarely because they entailed openly admitting lack of comprehension in the middle of an oral proficiency test. Of the eight students, only Pedro, Alejandro and Ying had listening comprehension problems, and only Pedro and Alejandro admitted that they did not understand a question or instructions. Ying used the strategy of silence, followed by an inappropriately cued confirmation check when he did not understand. Neither Alejandro nor Pedro understood the instruction "pretend that you are meeting your class on the first day to explain basic course requirements."

[12] Eval: (Reads the instructions as written about presenting the syllabus)

Pedro: Sorry I didn't catch you.

Eval: OK.

Pedro: I didn't understand you.

Eval: OK what I want you to do, is act like you are the teacher, and we are the class,

Pedro: OK

Eval: and act like, you know, you are giving this syllabus to us, don't really read it off the paper, but just look at it for a minute, and then, present the information to us, and this is your first day of class.

Pedro: Well, if the class is in, is in the afternoon, I would say, good afternoon, my name is Pedro, eh ... I'm your.. I'm the teaching assistant in charge of this course. I.. my.. I have my office in ... eh what, whatever, I'm not.. you can go to my office, Tuesday and Thursday,

Pedro's admission of not understanding the syllabus instructions in section four was tentative and nervous. He was only able to role play after hearing instructions that were more conversational in structure. Alejandro, after considerable negotiation using confirmation checks and clarification requests, began to talk about the syllabus in the third person The evaluators continued to simplify the instructions, but Alejandro never adopted a first-person instructor persona. His nonverbal behavior continued to be very confident as he turned the task into a conversation of mutual contingency. One evaluator actually blamed the role-play instructions rather than Alejandro's comprehension difficulties for the communication breakdown.

Alejandro also did not understand the instructions in example 4 and said so. His nonverbal behavior, however, managed to give the impression that the fault was the evaluator's rather than his own. As he listened to the instructions he leaned forward, bringing his ear as close as possible while maintaining eye contact. After a short time he began to shake his head steadily, smiling at the evaluator. Typically, such behavior can signal "you're wrong" or "I don't agree with you," and this was the message Alejandro conveyed to the evaluators. His demeanor was very confident, even though his subsequent response (example 4) indicated that he had indeed failed to comprehend.

The range of clarification requests exhibited in the data indicate that oral proficiency interview test takers can negotiate in a way that changes the power dynamic of the interview. By making use of features of conversational interaction, they can convince the evaluators that they have a higher level of proficiency than may in fact be the case from a purely linguistic perspective. The strategy worked for the interactionally competent students, Shiru and Alejandro, who were less linguistically proficient.

Framing the Discourse

Once the candidate's had taken the floor, they had to manage their discourse. In this study, successful interviewees framed the discourse to create interpersonal involvement and expectations about content knowledge. This finding supports the claims of Goffman (1974, 1981) and Tannen (1993) that participants in an interaction essentially define their own role and that of their interlocutors, and that these frames determine how the discourse is interpreted.

In an oral proficiency test in which only the first pre-planned section is in the candidate's discourse domain, interviewee's have to call on their knowledge of content and discourse structure in order to succeed. Successful test takers created appropriate expectations in their listeners as to their level of expertise in a given topic. Their obvious interpersonal involvement in their topics also created rapport. Each will be considered in turn.

Framing for content knowledge expertise

Example 13 below, illustrates how Shiru successfully framed her discourse to inform the evaluators at the outset that they could not expect much expertise on the topic.

[13] Eval: Describe the national sport in China, in as much detail as possible.

Shiru: National sports? [Possibly received a confirmatory nod - not visible.] OK. um actually I'm not very interested in sports, (laughs) um but I know something about that,

The discourse that followed was hesitant and halting. An evaluator reported "I remember at this part wondering if she quite understood or if she was just not into sports at all and couldn't come up with anything." She went on to say that she felt her hesitancy was "a thinking stumble"

because she lacked a schema to discuss sports in detail. In example 14, Lu's repetition, with the appropriate falling intonation, both acknowledged the evaluator's question, and gave her the planning time she needed to equate the notions of "teacher" and "designer." She framed the discourse within her own domain of computer assisted instruction, thus ensuring that the evaluators would not judge her off-topic.

[14] Eval: Question three, what do you think, are the most important characteristics, of an effective teacher?

Lu: Teacher, yeah, em .. for teacher and designer, for instance for me, I think the problem is, how can I provide, the best way for my students, to get the information.

Framing to create interpersonal involvement

Competence in interactional behavior created involvement between the evaluators and the candidate that ensured high scores for linguistically competent students, and passing scores for linguistically weaker interviewees. Two strategies will be singled out for discussion here, bonding moves, and disclosing personal information.

First, and perhaps most important, were bonding moves (Bailey, 1982) made by the candidates. Those who framed the interview as an opportunity to negotiate and reduce the unequal power dimension were more successful than those who accepted and maintained the asymmetry. Bonding moves include interactive listening behavior and appropriate turn-entry devices, but more than this, they include a willingness to step outside the immediate context and give listener-based responses. Ana did this when she compared her hometown affectionately with her U.S. university town . Two similar examples in the data demonstrate the difference between a bonding move and a mechanical move. Wang and Shiru received follow-up questions about attending a wedding after their presentation of the syllabus, as examples 15 and 16 show:

[15] Shiru: There will be no make up exams, um .. except for the medical reasons, um .. the syllabus uh.. will be, modified later, uh, any questions?

Eval: I'm getting married on October 20th, and that's when we have our first exam,

it's not for medical reasons, but can we work out a deal?

Shiru: Um ... [3 second pause] I think I will give you the exception, (laughs)

Eval: (laughs)

Shiru: because it's the, it's the most important moment, in your life, when you get married. Congratulations?! (laughs)

Eval: Thank you. If only it were true... (laughs) OK thank you.

Shiru: Thank 'you.

After Wang had explained his make-up policy, he was asked a similar question in example 16.

[16] Eval: I have two friends, that are getting married this quarter, and it, I would like to attend their wedding, it happens to be on the date of the exam, is that, a good enough excuse to be excused from the exam?

Wang: Um .. no this really is not a good reason, for you to ask for a, make up examination, you know, you know we have this every day at school, we have so many students, they have, they all have, some very close friends, so, so .. eh really I'm sorry I can't, allow you, to make up examination.

Eval: OK. Any other questions? OK. Thank you very much.

In Shiru's case, the evaluator said that she herself was getting married, and in Wang's case, the role-playing evaluator wanted to go to friends' weddings. Most instructors would immediately recognize a difference in kind in these two requests, and would probably feel a request by a student for "a deal" to enable her to attend her own wedding a reasonable one, whereas the request to go to friends' weddings would probably not evoke the same urgency. The main difference in the evaluators' reactions to the "yes" and "no" responses was that Shiru created rapport and interpersonal involvement, while Wang did not.

When Shiru asked if there were any questions, she put the syllabus down and placed her hands on the table, leaning forward towards the evaluators. On hearing the question, her eyes widened and her mouth opened in surprise. She looked down at the syllabus as though seeking the answer there, then looked up, tilted her head, smiled, and made her response. Her tone

conveyed genuine warmth. After the laughter ceased, Shiru continued to interact with correct stress placement on "you" in response to the evaluator's formal "thank you."

In contrast, Wang made no change in his discourse style between his monologue and what should have been an interactive response. His role play with the syllabus was hesitant, with uneven rhythm and pausing. He lacked authority, or teacher-voice (a quality an evaluator attributed to Lu). Wang gave his response to the question in the same narrow pitch range, with the same lack of animation, as he gave his presentation on the basic syllabus requirements. There was no enthusiasm and no attempt at bonding. The interaction remained within an interview frame, as is illustrated by the fact that it was the evaluator, not Wang, who closed the role play by asking other evaluators if they wished to ask any more questions.

Shiru also created involvement through her disclosure of personal information, as the following example shows:

[17] Shiru: Chinese, lot of Chinese, er choose their spouses, considering the economic factors,

Eval.: OK

Shiru: I chose my husband because we, we have a lot in common, and ... em I think we love each other, that's the important thing, the most important (laughs)

Eval: That's nice to hear. OK

Shiru: Thank you.

In this exchange, Shiru revealed personal information and received a rather dismissive interview frame response in return. She persevered in the interaction, however, treating it as a meaningful response, as she did consistently throughout her interview. This tactic eventually succeeded in drawing genuine responses from the evaluators. This finding supports those of Nelson (1991), who found that self disclosure was an effective strategy for establishing friendly interpersonal relations in the classroom, leading to higher evaluations from undergraduates.

Discussion

The results of the analysis highlight the complexity of face-to-face oral proficiency interview tests. The speech event in question was a gatekeeping test to determine whether the ITAs were sufficiently comprehensible to be assigned classroom related instructional duties. Thus, as several critics have noted about such tests (Lantolf & Frawley, 1988; Young & Milanovic, 1992; van Lier, 1989), it was indeed a situation where the power dynamic favored the evaluators who controlled the topics, and who expected the ITAs to "display." The evaluators interpreted the discourse management strategies and contextualization cues that were co-constructed in the interview process, and the more the test takers were able to approximate the culturally determined discourse and conversational management strategies expected by the native speakers of English, the higher they were evaluated.

Successful test takers in this study were able to modify the power dynamic and to reduce the interview asymmetry. In Goffman's terms, those who framed the interview as a discussion or conversation among peers, were more successful than those who framed it as an examination. The evaluators responded to the candidate's framing of the context. In the case of Alejandro, his framing of the context as a negotiated conversation in which he expected accommodation and assistance from the evaluators, enabled him to demonstrate pragmatic skills that resulted in his passing the test in spite of his poor listening comprehension and minimal fluency. As in Neu's (1990) study, Alejandro "bluffed" his way through.

In contrast, Carmen and Wang framed the interview as an examination and consequently did not perceive the need to establish rapport with their interviewers. Their only concern was to respond to the questions and complete the tasks as efficiently as possible. The evaluators read the directions and questions and did not interact with the candidate because they did not identify an interactive intent. The evaluators did not become interactively involved, and this incompatibility of cues was detrimental to the students' evaluations, as Erickson and Schultz (1982) demonstrated for counseling interviews. This analysis has shown that evaluators tend to spend less time and energy in the interview interaction when they are not engaged. As Gumperz (1982; 1992) and Tannen (1989) predicted, the examinees needed to demonstrate interactive

involvement with the evaluators to score well on the test, even though fundamental linguistic skills were good, as with Carmen. In their written comments, the evaluators commented on Carmen's lack of interactive skill, and expressed their doubts about her ability to do well as a TA in class. One evaluator commented, "the walls could have been asking her the questions and she would respond just the same." We claim that her low passing score of 3 is directly attributable to this perceived lack of interactional involvement. Similarly Wang did not succeed in involving the evaluators. He did not use many contextualization cues indicating supportive listenership and involvement, and, because of his intermediate or borderline linguistic skill, his final score was not improved by his low interactional competence.

Ying's linguistic competence was too low for him to compensate with the pragmatic skills demonstrated by Alejandro. He was below the threshold level posited in previous studies that have examined various contributing factors in oral proficiency (Bailey, 1982; Halleck & Moder, 1995). Thus, it seems that pragmatic competence assumes importance when linguistic competence is limited, as Seliger (1985) claimed. This study demonstrates that proficiency in discourse and conversational management strategies are critical for borderline test takers. Linguistically competent students pass regardless of interactional competence, but linguistically weaker examinees who employ strategies that enable them to negotiate meaning and sustain a conversation in a manner that engages the evaluators' interest and attention are also successful. They can transform a closed system into the authentic, dynamic open system required by Lantolf and Frawley (1988), and achieve the mutual contingency desired by van Lier (1989).

Implications

One of the major criticisms of oral proficiency interviews is that the unique roles and meaning created during the course of an interaction seriously compromise reliability. Each interview is unique because situated interaction implies variability. Shohamy (1988) included factors such as personalities and gender as variables that make the evaluation of oral language "complex, varied, and difficult to control " (p. 167). In the present study, the personalities of the ITAs' may have influenced their approach and responses to the interview format in subtle and

unknown ways. Kasper and Schmidt (1996) speculate that personality may be more of a factor in pragmatic rather than linguistic development. Certainly the role of personality features and gender characteristics in oral proficiency assessments must be investigated further.

Even though evaluators ask similar questions, and test takers perform similar tasks, the interactional construction of meaning will be different each time. We argue that this variability is also the potential strength of the oral proficiency test because it extends the notion of interactional authenticity (Bachman, 1990, 1991) to include the role of the evaluator as interlocutor. Discourse and conversational management strategies must be appropriate to the task at hand and can be most effective when they involve the evaluators on a personal level. Communication is not a mathematical equation, but the product of on-going interaction.

In acknowledging this fact, it is imperative to train both ITAs and evaluators to understand the role of contextualization cues in interpreting communicative intent. Discourse analysis studies such as this that look at more than linguistic competence should inform the classroom instruction of ITAs, test developers, and evaluator training. We need further research into inter-cultural and linguistic cues of, for example, the role of silence (Chafe, 1985; Scollon, 1985) or of the universality of acknowledging tokens (Clancy, Thompson, Suzuki, & Tao (1996). Although it is by no means certain that conversational management cues are amenable to classroom instruction (House, 1996), we must find ways to give students the opportunity to practice such skills as using comprehension checks, turn-taking devices that provide planning time (and understanding that native speakers expect initial disfluency while the discourse is being planned), learning appropriate nonverbal interactive behavior and good listenership strategies, encouraging or inviting questions from the audience, and framing their audiences expectations. It is also vital that test evaluators are made aware of the need to assess such interactional strategies appropriately so that they will not be seduced into ignoring serious linguistic difficulties. Similarly, because the interplay of communicative and linguistic factors in oral proficiency interviews is so complex, we must ensure that adequate reliability checks, balances and appeals processes are developed to protect the students.

References

American Council on the Teaching of Foreign Languages. (1986). ACTFL proficiency guidelines. Hastings-on-Hudson, New York: Author.

Bachman, L.F. (1988). Problems in examining the validity of the ACTFL Oral Proficiency Interview. Studies in Second Language Acquisition, 10, 149-164.

Bachman, L.F. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Bachman, L.F. (1991). What does language testing have to offer? TESOL Quarterly, 25, 671-704.

Bachman, L.F., & Palmer, A. S. (1983). The construct validity of the FSI oral interview, In. J.W.Oller, Jr. (Ed.). Issues in language testing research (pp. 154-169). Rowley, MA: Newbury House.

Bachman, L. F., & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. Modern Language Journal, 70, 380-390.

Bailey, K. M. (1982). Teaching in a second language: The communicative competence of non-native speaking assistants. Dissertation Abstracts International, 43, 3305A. (University Microfilms No. 83-06, 009.

Bailey, K. M. (1984). The "foreign TA problem." In K. M. Bailey, F. Pialorsi, & J. Zukowski/Faust (Eds.). Foreign teaching assistants in U.S. universities, (pp. 3-15).Washington, D.C: National Association for Foreign Student Affairs

Berwick, R, & Ross, S. (1993, August). Cross-cultural pragmatics in oral proficiency interview strategies. Paper presented at the Annual Language Testing Research Colloquium. Cambridge, England. (ERIC Document Reproduction Service No. ED 366 173)

Brown, K., Fishman, P., & Jones, N. (1990). Legal and policy issues in the language proficiency assessment of international teaching assistants (Monograph No. 90-1). University of Houston Law Center, Institute for Higher Education Law and Governance.

Chafe, W. (1985). Some reasons for hesitating. In D. Tannen & M. Saville-Troike (Eds.). Perspectives on Silence, (pp. 77-89). Norwood, NJ: Ablex.

Clancy, P. M., Thompson, S. A., Suzuki, R., & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. Journal of Pragmatics, 26, 355-387.

Douglas, D., & Selinker, L. (1985). Principles for language tests within the 'discourse domains' theory of interlanguage: test construction and interpretation. Language Testing, 2, 205-226.

Educational Testing Service. (1990). Test of Spoken English: Manual for score users. Princeton, NJ: Author.

Ericsson, K.A., & Simon, H.A. (1993). Protocol analysis: Verbal reports as data. Revised Edition. Cambridge, MA: MIT Press.

Erickson, F., & Schultz, J. (1982). The counselor as gatekeeper: Social interaction in interviews. New York: Academic Press.

Goffman, E. (1974). Frame analysis New York: Harper & Row.

Goffman, E. (1981). Forms of talk. Philadelphia: University of Pennsylvania Press.

Grice, P. (1975). Logic and conversation. In P.Cole, & J. Morgan (Eds.). Syntax and semantics, vol.3. New York: Academic Press.

Gumperz, J. J. (1982). Discourse strategies. Cambridge: Cambridge University Press.

Gumperz, J. J. (1992). Contextualization and understanding. In A. Duranti, & C. Goodwin (Eds.). Rethinking context. Cambridge: Cambridge University Press. 229-252.

Halleck, G.B. (1995). Assessing oral proficiency: A comparison of holistic and objective measures. The Modern Language Journal, 79, 223-234.

Halleck, G.B. (1992). The oral proficiency interview: Discrete point test or a measure of communicative language ability? Foreign Language Annals, 25, 227-231.

Halleck, G.B., & Moder, C.A. (1995). Testing language and teaching skills of international teaching assistants: The limits of compensatory strategies. TESOL Quarterly, 29, 733-758.

Hoekje, B., & Linnell, K. (1994). "Authenticity" in language testing: Evaluating spoken language tests for international teaching assistants. TESOL Quarterly, 28, 103-126.

House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. Studies in Second Language Acquisition, 18, 225-252.

Author, 1996 [Not entered to ensure anonymity].

Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. Studies in Second Language Acquisition, 18, 149-169.

Inglis, M. (1993). The communicator style measure applied to nonnative speaking teaching assistants. International Journal of Intercultural Relations, 17, 89-105.

Lantolf, J., & Frawley, W. (1988). Proficiency: Understanding the construct. Studies in Second Language Acquisition, 10, 181-195.

Madden, C.G., & Myers, C.L. (Eds.). (1994). Discourse and performance of international teaching assistants. Alexandria, VA: TESOL.

Nelson, G.L. (1991). Effective teaching behavior for international teaching assistants. In J. Nyquist, R. Abbott, D. Wulff, & J. Sprague (Eds.), Preparing the professoriate of tomorrow to teach (pp. 427-434). Dubuque, IA: Kendall/Hunt.

Neu, J. (1990). Assessing the role of nonverbal communication in the acquisition of communicative competence in L2. In R.C. Scarcella, E.S. Andersen, & S.D. Krashen (Eds.). Developing communicative competence in a second language. New York: Newbury House. 121-138.

Ross, S. (1992). Accommodative questions in oral proficiency interviews. Language Testing, 9, 173-185.

Ross, S., & Berwick, R. (1992). The discourse of accommodation in oral proficiency interviews. Studies in Second Language Acquisition, 14, 159-176.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. Language, 50, 696-735.

Scollon, R. (1985). The machine stops: Silence in the metaphor of malfunction. In D. Tannen, & M. Saville-Troike (Eds.). Perspectives on silence. Norwood, NJ: Ablex. 21-30.

Seliger, H.W. (1985). Testing authentic language: The problem of meaning. Language Testing, 2, 1-15.

Shohamy, E. (1988). A proposed framework for testing the oral language of second/foreign language learners. Studies in Second Language Acquisition, 10, 165-179.

Tannen, D. (1979). What's in a frame? In R. Freedle (Ed.)., New Directions in Discourse Processing. Norwood, NJ: Ablex. 137-181.

Tannen, D. (1989). Talking voices: repetition, dialogue, and imagery in conversational discourse. Cambridge: Cambridge University Press.

Tannen, D. (1993). Framing in discourse. Oxford: Oxford University Press.

Thomas, C. F., & Monoson, P. (1993). Oral English language proficiency of ITAs: policy, implementation, and contributing factors. Innovative Higher Education, 17, 195-209. Language Acquisition, 14, 71-86.

van Lier, L. (1989). Reeling, writhing, drawling, stretching, and fainting in coils: Oral proficiency interviews as conversation. TESOL Quarterly, 23, 489-508

Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. Studies in Second Language Acquisition, 14, 403-424.

Yngve, V. (1970). On getting a word in edgewise. In M. A. Campbell et al., Papers from the Sixth Regional Meeting of the Chicago Linguistic Society. Chicago: Chicago Linguistic Society. 567-578.

Appendix A

Oral English Proficiency Test

Preparing for the Test: Sample Questions

Warm-up: The evaluators will ask a few simple questions about yourself at the beginning of the test. This section is NOT scored. It will give you a chance to see that the evaluators are friendly people, and give you time to relax a little.

Section 1: Define a basic concept that you might teach to beginning undergraduates in your field of study. [Time: 3 minutes, including questions]. Choose a term or concept before you come to the test. You will have two minutes to explain your definition and 1 minute to answer questions from the evaluators. You may use the blackboard, and may bring only one 3 x 5 inch index card containing a brief outline to remind you of your main points. We will evaluate you on your ability to explain the concept clearly and simply, as though we were your students. Keep it simple and clear, and don't memorize your speech.

Examples:

Math: Derivative

Mechanical Engineering: Force

Chemistry: Molecule

Physics: Inertia

English: Adjective

Section 2: We will ask you three questions asking you to describe or give your opinion about events, objects or concepts of general interest. There are no right or wrong answers to these questions. You should respond freely and in as much detail as possible. Think of this as a conversation.

Examples:

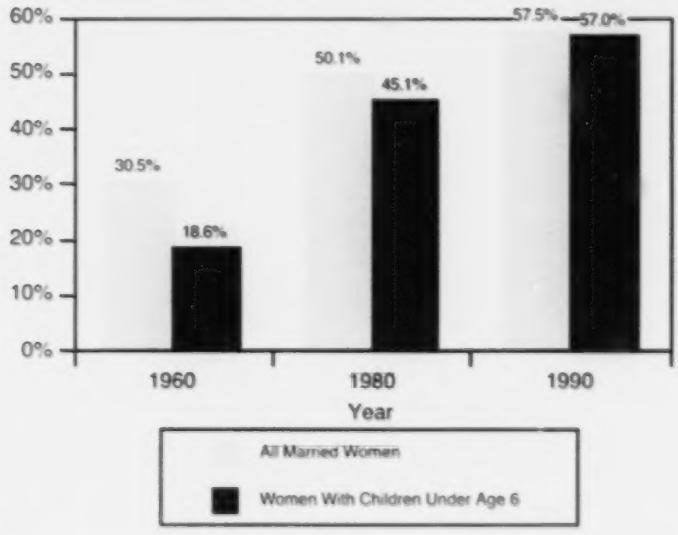
1. Tell us about your hometown.
2. Tell us how people in your country prepare themselves for their career.
3. What do you think are the most serious environmental problems facing us today?

Section 3: In this section we will ask you to explain information about a general topic presented in a chart or figure. You will have one minute to study the information. When you have studied the graph, you will be asked to go to the overhead projector and to explain the most interesting information in the figure or chart to the evaluators [Time: 1 minute]. We do not expect you to cover all the material in the short time available. The evaluators will ask you questions about your explanation.

If you have never used the Overhead Projector before, you may come early to the test so that you can practice. But we are NOT testing your ability to use the Overhead Projector.

Example:

Percentage of U.S. Women in the Workforce



Section 4 : In this section we will ask you to assume the role of an instructor on the first day of class. Pretend that you are talking to a group of students. We will provide you with a short syllabus and after studying it for a minute, you may go to the Overhead projector and explain the information to your students (role-played by the evaluators). The evaluators will ask you

questions at the end of your explanation. It is important to use your own words in talking about the information. Do not simply read the information aloud.

Example:

COURSE REQUIREMENTS

Readings:	Reading assignments to be completed before each class.
Homework:	Weekly assignments due Friday at beginning of class.
Exams:	Midterm: Wednesday of week 5
	Final: See Exam week schedule.
	NO MAKE-UP EXAMS EXCEPT FOR MEDICAL REASONS
Attendance:	Required. More than two absences requires you to drop the course.

Appendix B

Oral English Proficiency Test Scoring Key

Pronunciation

1. Poorly pronounced individual sounds, poor articulation, excessively fast rate, foreign rhythmic patterns. Unintelligible.
2. Serious difficulties with individual sounds, articulation, excessively fast rate, rhythmic patterns. Often unintelligible.
3. Generally accurate pronunciation of individual sounds, articulation, appropriate rate of speech and rhythmic patterns. Minor problems only. Generally intelligible at all times.
4. Native-like pronunciation of individual sounds, articulation, rhythmic patterns and appropriate rate of speech. Always intelligible.

Grammar

1. Can express only a few memorized phrases accurately. Has difficulty in producing even the simplest utterances grammatically and appropriately.
2. Frequent errors in grammatical forms and sentence structures. Structures are too basic, or errors are serious enough to interfere with the meaning the speaker is trying to express.
3. Appropriate control of most grammatical forms and sentence structures. Minor errors do not interfere with the meaning the speaker is trying to express.
4. A wide range of grammatical forms and sentence structures that are almost always accurate and appropriate. Occasional errors that might be made by a native speaker.

Fluency

1. Does not understand questions, even after repetition; speech is hesitant, fragmentary, and is limited to a few simple words and phrases.
2. Frequently misunderstands and needs repetition, or slowed down and rephrased speech. Cannot clarify meaning by negotiation with the evaluator. Lack of vocabulary may cause non-native pauses and hesitations (may be translating). Difficulty with smooth rhythmic patterns causes frequent intelligibility.
3. Able to understand most audience questions and comments without difficulty. Can negotiate meaning successfully. Vocabulary is generally accurate and appropriately used. Occasional non-native pauses and rhythmic patterns do not interfere with intelligibility.
4. Able to understand speech at a natural pace. Does not require extensive adjustments. Vocabulary is accurate, idiomatic and appropriate. Very few non-native pauses. Rhythmic patterns are smooth and native-like.

Comprehensibility

1. Speech is basically unintelligible. Major errors in pronunciation, grammar, and/or fluency make communication impossible.
2. Speech is frequently incomprehensible. Significant effort and concentration are required by the listener to understand the message. Not ready for teaching.
3. Speech is generally comprehensible. Some errors in pronunciation, grammar, and/or fluency do not interfere with communication. Excessive concentration is not required to understand the message. Ready for teaching.
4. Completely comprehensible, native-like effective communication. Minor errors only which do not interfere.

Appendix C

Transcription Conventions

(from Tannen, 1989)

.	sentence final falling intonation.
,	clause-final intonation ("more to come").
?!	exclamatory intonation.
...	pause of 1/2 second or more.
..	perceptible pause of less than 1/2 seconds.
CAPS	emphatic stress.
:	following a vowel, indicates elongation.
::	extra elongation.

Table 1: Participants' Gender, Country, Department, and OEPT Score for Overall Comprehensibility.

Participant/ Gender	Country	Department	OEPT Score (Scale 1-4)
Ana/F	Peru	Spanish	3.9
Alejandro/M	Venezuela	Music	3.1
Carmen/F	Puerto Rico	Spanish	3.0
Pedro/M	Venezuela	Physics	2.6
Lu/F	PRC	Education	3.7
Shiru/F	PRC	Neuroscience	3.0
Wang/M	PRC	Chemical Eng.	2.8
Ying/M	PRC	Math.	2.3